DOCUMENT RESUME

ED 376 422 CG 025 843

AUTHOR Hummel, Thomas J.
TITLE Probabilities of Client Success in Counseling.
PUB DATE 94
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 4-8, 1994).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Counseling; *Counseling Effectiveness; *Effect Size; Expectation; Mathematical Models; *Outcomes of Treatment; *Predictor Variables; Probability; Research Methodology; Statistical Significance; *Statistical Surveys; Therapy

ABSTRACT

Researchers should investigate statistical models that can help counselors decide how to treat individual clients. This research investigated the questions, "Given an effect size (ES) from a counseling outcome study, what is the probability that a client would have a negative response to the treatment, and what is the probability that the client would receive a minimally acceptable benefit?" Statistical models of 144 experimental situations were evaluated. ES, sample size, the variance of clients' treatment reactions and the distribution of these reactions were varied. With respect to the probabilities of a negative response to treatment, it was found that if the counselor is willing to assume that all clients respond identically to a treatment, then the counselor need only attend to the sample sizes and the effect size. If the effect size is a least 0.50 or the sample size is at least 15 in each group, the probability is less than 0.10 that the true effect size is negative. This percentage increases as the effect size decreases and the standard deviation of the individual responses to treatment increases, regardless of the sample size. Similar patterns hold for determining the probability of effect sizes exceeding a minimally acceptable benefit. Two tables provide statistical summaries. Contains eight references. (RJM)

Probabilities of Client Success in Counseling

By

Thomas J. Hummel
University of Minnesota

A Paper Presented at the Annual Meeting of the
American Educational Research Association
New Orleans, 1994

Abstract

This research investigated the questions, "Given an effect size (ES) from a counseling outcome study, what is the probability that a client would have a negative response to the treatment, and what is the probability that the client would receive a minimally acceptable benefit?" Statistical models of 144 experimental situations were evaluated, and in 135 of these, clients could have a unique, individual reaction to treatment. ES, sample size, the variance of clients' treatment reactions and the distribution of these reactions were varied. ES was taken to be the fixed value of a random variable, and given this value, the likelihood of various values of the true, population effect size were computed. For a given experimental situation, a probability was computed by first determining the conditional probability of a client reaction given a population effect size, and then, allowing the population effect size to vary throughout its range, finding the weighted sum of the conditional probabilities using the likelihoods of the population effect sizes as the coefficients. Probabilities were computed using numerical integration routines provided by Mathcad 4.0. The major findings were: 1) as the variability of client responses to treatment increases, so does the probability that a client will have a negative response to treatment, and 2) sample size has little effect on the probability that a client will have a negative response to treatment.

Introduction

Counselors do research for a great variety of reasons. One goal is to better understand people and their individual differences. In this vein, their work is sometimes hard to distinguish from that of personality researchers and social psychologists. Another goal is to evaluate programs to determine if their effectiveness warrants the resources committed to them. For example, does a particular drug prevention program in a high school lower the incidence of drug abuse? A third goal is to research a treatment or intervention designed to help individual clients with some problem. It is this last type of research that is the focus here. For brevity, it will be referred to as outcome research.

This symposium deals with the question, "Can counselors use outcome research to justify the treatment of a particular client?" This question is different from the question, "Can counselors use outcome research to justify the counseling of clients?" The latter suggests an evaluation of the profession as a whole, or major segments of it, and is not of primary concern here. The present focus is on the individual counseling relationship and how useful counseling research is in helping a counselor decide whether a client will benefit from a particular intervention. Specifically, given a particular treatment, what is the probability that a client will benefit from counseling?

If the linear models taught to counselors in their statistics courses actually obtained, justifying the treatment of individual clients would be straightforward. In the simple model for a one-way design, $X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, it is assumed that all clients have exactly the same response to the treatment. The effect size for the group is the effect size for each client.

Based on an informal survey of counselor trainers and trainees, no one seems to believe that all clients receive exactly the same benefit from a treatment. In fact, the consensus seems to be that while some clients might thrive, others might actually get worse. This means that the model $X_{ij} = \mu + \alpha_{ij} + \varepsilon_{ij}$ probably better represents counselors' beliefs, for in this model each client can have his or her own unique treatment response. In a program evaluation, one could consider the model $X_{ij} = \mu + \overline{\alpha}_{i.} + \varepsilon'_{ij}$. Here one considers the average treatment response and includes the individual's deviation from that average with error, or $\varepsilon'_{ij} = \alpha_{ij} - \overline{\alpha}_{i.} + \varepsilon_{ij}$. While this might satisfy a funding agency, it would be little comfort to a client or counselor, for the client may be one of the unlucky ones who deteriorates with such a treatment.

The counseling profession needs to investigate statistical models that can better help counselors decide how to treat individual clients. What happens on the average in a group is important, but it cannot take the place of an assessment of what happens to individuals. If a counselor knew that on the average clients *improved* with a given treatment, but also knew that, say, 55% of clients actually *got worse*, would the counselor be justified in using the treatment? This situation could arise if individual client treatment responses were positively skewed. In this case, a counselor would probably not use the treatment, for it

would be more likely than not that the client would deteriorate. However, the counselor's decision would depend on a consideration of more than the effect size (ES) associated with the treatment. The distribution of the "individual effect sizes" (IESs) of the clients would need to be considered. Presently, the analysis of the typical outcome study and its reporting does not consider the distribution of individual effect sizes.

Ultimately, the goal is to relate research findings more directly to the optimal treatment of individual clients. One way to think about this problem is to conceptualize a counselor who wishes to use the results from a treatment-control outcome study to determine the course of treatment for a client. This counselor decides that the client may reasonably be considered a member of the population studied and wants to use the obtained effect size (ES) in determining how likely it is that the client will benefit. To determine benefit, individual effect sizes (IESs) are considered, i.e., an IES represents each client's individual response to the treatment. The counselor is interested in two probabilities: the probability that a client would have a negative response to the treatment, $Pr[IES < 0]$, and the probability that the client's treatment response would exceed a "minimal acceptable benefit (MAB)," or $Pr[IES \geq MAB]$. The MAB is the smallest positive treatment outcome which justifies the client's investment of time and money. There are, therefore, positive IESs less than MAB which are too small to justify treatment.

It is these two probabilities, $Pr[IES < 0]$ and $Pr[IES \geq MAB]$, that are the focus of this paper. The next section presents the models used to evaluate these probabilities.

Related Literature

Other authors have described models that bear some similarity to the one considered here. Two are discussed briefly. Wood and Games (1990) have written an expository paper on a specific class of underspecified linear models. They consider, as one example, "differential treatment effects across unmeasured subpopulations." Later in the present paper, it will be seen that one of the models studied is related to this example of Wood and Games's. Their approach, however, is to develop a model of unmeasured interactions with the constraint that the correlations between unmeasured and measured independent variables be positive. Further, they describe their model as appropriate for describing differential rates of improvement, and do not include a situation where a client might actually deteriorate as function of treatment. The approach taken here does not assume the conditions that they do.

Hedges and Olkin's (1985) book on "Statistical Methods for Meta-Analysis" includes a chapter on "Random Effects Models for Effect Sizes." These models are motivated by the realization that large scale evaluation studies involving multiple sites may in effect be assessing treatments that are unique to the site in certain respects. If one imagines a client being an individual "site," then Hedges and Olkin's approach has some similarity to that taken here. However, their assumption of within site consistency of treatment effect allows them to avoid the confounding of experimental error and individual response to treatment that is inherent in the present approach, i.e., there is no directly analogous way

to compute an effect size for an individual client. Also, whether their focus is on the across site average effect size or the variability of site effect sizes, their model does not relate directly to the central concern here, namely, the probability of client success.

Probability Models

The general research situation modeled was that of a randomized, two-group experiment, with a single treatment and a no-treatment control group. The linear model introduced above to reflect individual treatment responses, $X_{ij} = \mu + \alpha_{ij} + \varepsilon_{ij}$, is modified slightly to reflect this experiment by changing the first subscript to "T" or "C." The treatment group's model is $X_{Tj} = \mu + \alpha_{Tj} + \varepsilon_{Tj}$, and for the no-treatment control group, the model is $X_{Cj} = \mu + \varepsilon_{Cj}$. Given these models, the population effect size is $\mu_{ES} = \frac{\overline{\alpha}_T}{\sigma}$, and it is estimated by $ES = \frac{\overline{X}_T - \overline{X}_C}{s}$. An individual effect size is defined as $IES = \frac{\alpha_{Tj}}{\sigma}$. (Note that the usual side condition that the treatment effects sum to zero is not imposed.)

When a counselor reviews a research paper and notes the effect size, ES, it would be appropriate to consider the probability that a client might deteriorate because of the treatment, i.e., the probability that an IES is less than zero. To do this, the distributions of ES and IESs must be considered. Further, two sources of variability, the standard deviation of the IESs, $\sigma_{IES} = \frac{\sigma_{\alpha_{Tj}}}{\sigma}$, and of ES, $\sigma_{ES} = \sqrt{\frac{2}{n}}$, must also be considered. (See Note 1.)

In the analysis that follows, the outcome variable was assumed to have a standard deviation of $\sigma = 1.00$. The sample size for each group, treatment and control, was equal to n and was alternatively set to n = 15, 35, and 90. The effect sizes investigated were set at ES = 0.3, 0.5, and 0.8. An ES = 0.8 would lead to a statistically significant result (α = .05) for all values of n studied. For ES = 0.5, significance would be achieved for n = 35 and 90, and for ES = 0.3, only n = 90 would lead to significance. Therefore, six of the nine combinations are statistically significant, and this seems reasonable because those reading the research literature would tend to ignore nonsigificant results. On the other hand, with the increased popularity of quantitative integrations of research, primarily meta analyses, it seemed prudent to also include several combinations that, while not significant, might play the role of rather weak treatments that are deemed to have an effect due to the consistency with which the effect appears across studies.

In order to study the effect of the IESs, it was necessary to consider both the variability and shape of the distribution of these quantities. Four levels of variability were investigated, σ_{IES} = 0.00, 1/3, 2/3, and 1.00. The first value is simply used to determine a reference point, for the probabilities for 0.00 are those that would obtain if the usual model were correct and every one had precisely the same response to the treatment. The other values can be considered relative to the standard deviation of the criterion. The largest value, 1.00, indicates that there is just as much variability in the individual

responses to the treatment as there is individual variability on the variable itself. For example, assuming normal distributions for the moment, if 68% of individuals lie between 40 and 60 on the MMPI-D Scale, then 68% of the clients' responses to treatment would lie in the interval ES ± 10. An analogous interpretation holds for σ_{IES} = 1/3 and 2/3.

While the normal distribution is a reasonable one to assume for the distribution of the means, there is no reason to assume that IESs would be distributed in this manner. To study the effect of the shape of the distribution of IESs, five distributions were used. All of the following distributions were transformed to have a mean of ES and a standard deviation of σ_{IES}. The distributions investigated were:

1. Normal Distribution

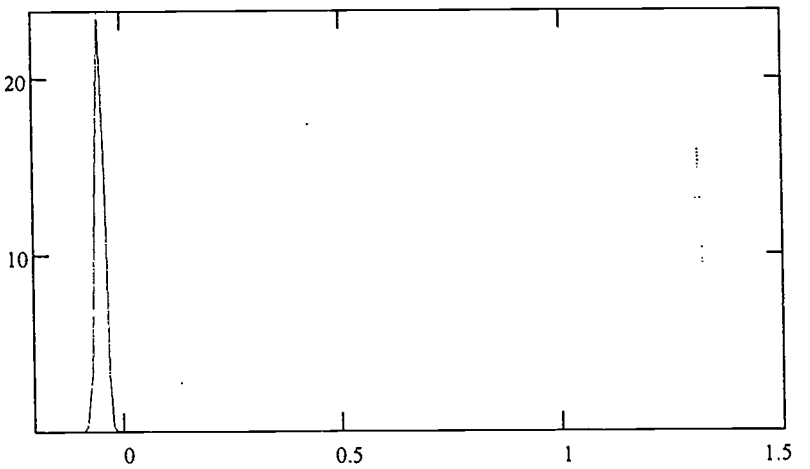
2. Uniform Distribution: The primary motivation for using this distribution was to include a symmetric distribution with larger tail area probabilities than the normal distribution. Considering distributions of IES where $ES = \mu_{ES}$ and $\sigma_{IES} > 0$, the probability of a negative outcome is equal to the area below zero. For the nine combinations of ES and $\sigma_{IES} > 0$ investigated, the transformed zero point varied between the standard scores $z = -2.4$ and $z = -0.3$. For $z = -2.4$ the normal distribution gives a greater area than the uniform. At $z = -1.5$, the areas are equal. For the remaining seven conditions studied, the uniform distribution gives the greater area. The area below $z = -1.2$ in the uniform distribution equals the maximum obtainable for any symmetric, unimodal distribution (Glass and Hopkins, 1984, p. 64).

3. Chi-Square Distribution on 4 Degrees of Freedom (Chi2 +): This distribution was included to investigate the effect of positive skewness ($\gamma_1 = 1.414$). A positively skewed distribution would result if a relatively small number of clients had an extremely positive response to the treatment.

4. Reflected Chi-Square Distribution on 4 Degrees of Freedom (Chi2 -): This distribution was included to investigate the effect of negative skewness ($\gamma_1 = -1.414$). A negatively skewed distribution would result if a relatively small number of clients had an extremely negative response to the treatment.

5. Mixture of Two Normal Distributions: The distributions studied were bimodal, the kind that would result if there were distinct subpopulations within a population under study. These distributions served as a "bad case," and they were specifically designed to increase the probability of negative responses to the treatment. For Subpopulation One, the mean and standard deviation were $\mu_1 = -5\sigma_1$ and $\sigma_1 = 0.01$, respectively, and its proportion in the mixture was Π_1. For Subpopulation Two, the corresponding parameters were $\mu_2 = \dfrac{ES - \Pi_1\mu_1}{\Pi_2}$ and $\sigma_2 = 0.01$, with $\Pi_2 = 1 - \Pi_2$. The definition of μ_1 results in essentially all of Subpopulation One being below zero (negative IESs),

while the definition of μ_2 guarantees that $ES = \Pi_1\mu_1 + \Pi_2\mu_2$. The value of Π_1 was chosen to satisfy the equality:
$\sigma_{IES} = \sqrt{\Pi_1\left[\sigma_1^2 + (\mu_1 - ES)^2\right] + \Pi_2\left[\sigma_2^2 + (\mu_2 - ES)^2\right]}$. Since the value of σ_{IES} depends on the distance between the subpopulations, as well as the variability within them, setting σ_1 and σ_2 to relatively small values allowed the subpopulations to be farther apart. As a result of the preceding definitions, the relative sizes of the subpopulations and the distance between them varied as a function of ES and σ_{IES}. As an example, when ES = 0.5 and σ_{IES} = 2/3, for Subpopulation One $\mu_1 = -0.05$, $\sigma_1 = 0.01$, and $\Pi_1 = 0.595$, while for Population Two, $\mu_2 = 1.308$, $\sigma_2 = 0.01$, and $\Pi_2 = 0.405$. These subpopulations are depicted in the following graph:



The mixture based on these two subpopulations is positively skewed ($\gamma_1 = 0.387$), but that would not always be the case. For values of ES and σ_{IES} studied, the skewness varies from -2.156 to 2.507 and was never equal to zero. Given the constraint that virtually all of Subpopulation One would be below zero, it follows that as $\frac{ES}{\sigma_{IES}}$ increases, Π_1 must decrease. When $\Pi_1 < 0.5$, the mixture is negatively skewed.

As mentioned previously in this section, all distributions were located at ES, the effect size. ES is a random variable and it would be unlikely that the true value, μ_{ES}, would equal ES. To account for sampling variation, the likelihood of various values of μ_{ES} given ES were computed. The normal density function was used as the likelihood function for μ_{ES} because it is easy to compute and would provide a close enough approximation to the appropriate noncentral t-distribution for current purposes. (See Note 2.) The likelihood for μ_{ES} was defined, therefore, as $\frac{1}{\sigma_{ES}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\mu_{ES} - ES}{\sigma_{ES}}\right)^2}$. Given the above setup, the probability that an IES would fall below zero was defined as:

$$Pr[IES < 0] = \frac{1}{\sigma_{ES}\sqrt{2\pi}} \int_L^U e^{-\frac{1}{2}\left(\frac{\mu_{ES} - ES}{\sigma_{ES}}\right)^2} \left[\int_{L'}^{0} f_{IES}(\mu_{ES})\, dIES\right] d\mu_{ES}.$$

The probability that IES would fall above the minimal acceptable benefit (MAB) was defined as:

$$\Pr[\mathrm{IES} \geq \mathrm{MAB}] = \frac{1}{\sigma_{ES}\sqrt{2\pi}} \int_{L}^{U} e^{-\frac{1}{2}\left(\frac{\mu_{ES}-ES}{\sigma_{ES}}\right)^2} \left[\int_{MAB}^{U'} f_{IES}(\mu_{ES}) \, dIES \right] d\mu_{ES}.$$

In the preceding integrals, the density of the IESs, $f_{IES}(\mu_{ES})$, was alternatively defined according to the five distributions defined earlier in this section. The parameter, μ_{ES}, caused the distribution to be relocated as μ_{ES} varied throughout its range in the outer integral. The limits of integration for the outer integral were $U = ES + 5\sigma_{ES}$ and $L = ES - 5\sigma_{ES}$. The limits for the inner integral varied depending on the density function, but it was generally deemed adequate to consider the probability content in a region $10\sigma_{IES}$ wide.

With respect to computing the probability of an IES being negative, the inner integral computed the conditional probability of an IES falling below zero, given that the distribution was located at μ_{ES}, while the outer integral obtained the weighted sum of these conditional probabilities, the weight corresponding to the likelihood of each value of μ_{ES} given ES. Therefore, conditional probabilities for more likely values of μ_{ES} were given a larger weight in the overall probability. The integrals for MAB have a similar interpretation.

All probabilities were obtained by numerical integration using Mathcad 4.0. The results obtained are included in the Appendix.

Findings

The numerical results are presented in Table 1 and Table 2 in the Appendix. Together, these tables contain 288 probabilities. The following description of the results includes attention to specific probabilities that seemed of interest, the identification of trends, and statistical analyses of the probabilities themselves. While an attempt has been made to report the salient features, readers are encouraged to study the tables to find relationships and points of interest that have not been included by the author.

The probability of a negative outcome

The first three columns of Table 1 give the values of n, ES, and σ_{IES} investigated. The next five columns give the probabilities obtained for the five distributions studied, Normal, Uniform, and so on. (The term "$\sigma_{IES} > 0$ columns" will be used to refer to this set of five columns.) The last column, Reference, gives the probability of μ_{ES} being less than zero when $\sigma_{IES} = 0$, i.e., when the treatment has precisely the same effect on all clients.

When $\sigma_{IES} = 0$, across the conditions studied, the Reference probabilities vary from 0.00 to 0.21. These are the probabilities that the treatment has a negative effect even though the observed effect size (ES) was positive. However, when σ_{IES} is allowed to assume the values $\sigma_{IES} = 0.33$, 0.67 and 1.00, the probability of a negative response to treatment ranges from 0.00 to 0.56. The highest probability in the Reference column and the highest values in the $\sigma_{IES} > 0$ columns do not occur under the same conditions. For example, 0.56 is found in the Mixture column when n = 90, ES = 0.30 and $\sigma_{IES} = 1.00$, while the highest value under Reference occurs when n = 15 and ES = 0.30. The conditions under which 0.56 occurs in the Mixture column result in a corresponding Reference value of only 0.02. This Mixture probability is over 27 times higher than the corresponding Reference probability. This discrepancy is due to the fact that sample size, n, has a relatively large effect on the Reference probabilities, but a much smaller effect on the $\sigma_{IES} > 0$ column probabilities. The correlation between the sample sizes and the Reference probabilities is -0.54, while the correlations between sample size and the probabilities in the $\sigma_{IES} > 0$ columns vary from only -0.13 to 0.11.

The Reference probabilities and $\sigma_{IES} > 0$ column probabilities are affected similarly by effect size (ES). The correlation between ES and the Reference probabilities is -0.63, while the corresponding probabilities for the $\sigma_{IES} > 0$ columns range from -0.74 to -0.61.

For the Reference probabilities, as sample size and effect size increase, the probability of a negative effect decreases. This, of course, is precisely what would be expected from theory. For the $\sigma_{IES} > 0$ columns, effect size has an important effect, but not sample size. The $\sigma_{IES} > 0$ columns, however, have an additional source of influence, the variability of the individual effect sizes (IESs). It is this source of variability that causes the $\sigma_{IES} > 0$ columns to have substantially different probabilities than those in the Reference column.

The $\sigma_{IES} > 0$ columns are highly related to one another. The correlations among these columns vary from 0.946 to 0.997. The first four columns, Normal through Chi2 -, have correlations from 0.980 to 0.997. The relationship of the Mixture column is only slightly less related to the other four, 0.946 to 0.977. This drop in correlation is likely due to the fact, as pointed out above, that the definition of the Mixture distribution varies as conditions change. Still, all the $\sigma_{IES} > 0$ columns are related highly enough to combine them while considering the influence of effect size (ES) and the variability of the individual effect sizes (IESs), as indexed by their standard deviation, σ_{IES}.

The multiple correlation between the row means of the $\sigma_{IES} > 0$ columns (i.e., the average obtained by summing across the columns in each row of Table 1 and dividing by five) and ES and σ_{IES} is 0.983. Since the conditions studied are uncorrelated, the individual contribution of ES and σ_{IES} equals the square of their correlations. For ES, it is $(-0.689)^2 = 0.475$, and for σ_{IES} it is $(0.702)^2 = 0.493$. (The multiple correlation is therefore $\sqrt{0.475 + 0.493} = 0.983$). Both ES and σ_{IES} have a strong influence on the row means. They each account for almost half the variance in this average, and indicate that

the probability of a negative reaction to treatment increases as ES decreases and σ_{IES} increases. While the direction of these relationships are what would be expected from a theoretical analysis of the model investigated, the extent to which a linear function of ES and σ_{IES} predicts the average probability was not anticipated.

To say that the $\sigma_{IES} > 0$ columns' probabilities are similarly influenced by the conditions studied is not to say that there are no differences among the columns. The column means for Normal through Mixture are 0.224, 0.251, 0.236, 0.193, and 0.321, respectively. The last two, Chi2 - and Mixture, are the more divergent, with Mixture's 0.321 being the most different. The manner in which the IESs are distributed does affect the probability of a negative response to treatment. As a point of comparison for these column means, the average of the Reference column is 0.050. The average of the column averages, i.e. the grand mean, is 0.245, nearly five times the size of the average of Reference probabilities.

From another perspective, the correlation ratio (η^2) for rows is 0.843, and for columns, it is 0.110. This means that over 95% of the variance in Table 1 is accounted for by row and column effects, leaving less than 5% due to the interaction between rows and columns. This indicates that the rows and columns have an additive effect on the probability that a client will have a negative response to treatment. This, along with the strong linear relationship that ES and σ_{IES} have with the row means, lead to a straightforward understanding of the influences on probabilities of a negative outcome.

Under certain conditions studied, substantial proportions of clients would be negatively affected by a treatment, even though that treatment was of benefit on the average. The probability of a negative outcome is an additive function of the conditions investigated, with ES, σ_{IES}, and the distribution of individual effect sizes playing a major role. Sample size had little influence on these probabilities. The smaller the ES, the greater the likelihood that the IESs will be negative. As σ_{IES} becomes larger, the distributions of IES spread out and more of their area falls below zero.

Sample size does affect the Reference probabilities, with the probabilities decreasing as n increases. The distribution of the μ_{ES} contracts, because it is less likely that the true value varies greatly from the sampled value as n increases. Since the probabilities obtained when $\sigma_{IES} > 0$ are not affected greatly by variations in sample size, the largest discrepancies between the Reference probabilities and the others are found when both n and σ_{IES} are large.

With respect to the conditions that lead to substantial probabilities of a negative response, one must ask, "How likely is it that they would be encountered in the real world?" This question is dealt with in the discussion section.

The probability of an "acceptable" outcome

While counselors should avoid hurting clients, they also need to strive to give clients some benefit that is commensurate with the time and money invested in therapy. To operationalize this goal, the term "minimal acceptable benefit (MAB)" is introduced. For this investigation, MAB was set at 0.5, meaning that a client would need to improve by one half standard deviation on the outcome measure in order to have received a minimal acceptable benefit (MAB). This criterion is arbitrary, and its only justification at this point is that it equals what is generally considered in the literature as a "moderate" effect size.

Table 2 presents the probabilities that an IES will exceed the MAB. These results contrast with those given in the preceding section in two important ways. First, in Table 1, *all* $\sigma_{IES} > 0$ columns' probabilities are higher than their Reference probability counterparts. In Table 2, however, 46% of the $\sigma_{IES} > 0$ columns' probabilities are less than their Reference probability counterparts, 13% are the same, and 41% are greater. In 41% of the cases, then, $\sigma_{IES} > 0$ leads to a larger probability that the client's benefit will exceed the MAB. When ES = MAB = 0.50, the symmetric distributions, the Normal and Uniform, have half their area on either side of the MAB, which, of course, agrees perfectly with the Reference probabilities. As for the other distributions, positive skewness is associated with probabilities being less than the Reference probabilities, for all of the Chi2 + probabilities are smaller, as are those for the instances where the Mixture is positively skewed. For ES = 0.80, and, therefore, ES > MAB, all probabilities in the $\sigma_{IES} > 0$ columns are less than the Reference probabilities, meaning the client is less likely to obtain the MAB when there are unique responses to treatment. When ES = 0.30, with three exceptions, the probabilities in the $\sigma_{IES} > 0$ columns are larger than the Reference probabilities. The three exceptions occur in the Mixture column where both sample size and $\frac{ES}{\sigma_{IES}}$ are smaller. The mixture is positively skewed in these three cases.

In Table 2, with very few exceptions, positive skewness is associated with lower probabilities than symmetry, while negative skewness is associated with higher probabilities. This makes conceptual sense since positive skewness results in more probability mass to the left of the mean and this tends to place more mass to the left of MAB. The opposite is true for negative skewness.

With some values in the $\sigma_{IES} > 0$ columns higher than the Reference probabilities and others lower, it is not surprising to find similar means for these sets of probabilities. The average Reference probability is 0.53, which is close to the average in the remainder of the table, 0.50. While the means are much more similar than in Table 1 (where the corresponding values were 0.05 and 0.25), the variance of the $\sigma_{IES} > 0$ columns' probabilities in Table 2 is 50% greater than in Table 1, and there are some marked discrepancies between the Reference probabilities and those in the $\sigma_{IES} > 0$ columns. For example, when n = 90, ES = .30, and $\sigma_{IES} = 1.00$, the Chi2 - probability of exceeding MAB is 0.51, while the corresponding Reference probability is 0.09. In this case, the individual treatment responses lead to five times as many clients reaching the minimal benefit as would be the case if the treatment response was the same for everyone. In the

other direction, when n = 90, ES = .80, and σ_{IES} = 1.00, the Mixture probability is 0.42 while the Reference probability is 0.98. The Reference probability would lead one to believe that only 2% of the time would ES fail to exceed the minimal benefit, but in actuality, 58%, the majority clients, would have an IES less than the MAB.

On a more global level, as in Table 1, sample size has very little influence on the $\sigma_{IES} > 0$ columns' probabilities, with correlations between sample size and these probabilities ranging from -0.04 to 0.06. With exception of the Mixture column, the probabilities in the $\sigma_{IES} > 0$ columns have a stronger linear relationship with ES than σ_{IES} (correlations from 0.91 to 0.94 for ES as opposed to correlations from -0.18 to 0.11 for σ_{IES}). It is clear in the first four columns of Table 2 that σ_{IES} has an effect. When ES = .30, increasing values of σ_{IES} are associated with increasing probabilities, but when ES = .80, increasing values of σ_{IES} are associated with decreasing probabilities. In the fifth column, Mixture, increasing values of σ_{IES} are associated with decreasing probabilities, regardless of the value of ES. This is in contrast with Table 1, where there is a similar linear influence for ES and σ_{IES}. The correlations with the Mixture column (0.74 and -0.64 for ES and σ_{IES}, respectively) are more in line with those for Table 1.

The first four $\sigma_{IES} > 0$ columns are highly related, with correlations ranging from 0.941 to 0.997. The correlations between Mixture and the other four distributions range from 0.66 to 0.83. These lower correlations are probably due to the fact that as the skewness of the Mixture varies from negative to positive values, the probabilities tend to decrease, thereby changing the ordinal position of these probabilities with respect to the other columns.

The means of the $\sigma_{IES} > 0$ columns vary from lows of 0.43 for Mixture and 0.46 for Chi2 + to 0.58 for Chi2 -. In terms of proportion of variance accounted for, "columns" accounts for 10% of the variance ($\eta^2 = 0.10$), which is much smaller than that accounted by the "rows" ($\eta^2 = 0.77$). In contrast to Table 1, interaction between rows and columns accounts for 13% of the variance in Table 2, which makes this source of variance larger than that for columns.

Discussion

This project was designed to give counselors some idea of the probability that a client would have a negative response to a treatment and, also, the probability of a client improving at least one half standard deviation on the treatment's outcome measure. It is assumed the counselor would have an effect size from a research paper and know the size of the samples studied. Taking this information from the research report and combining it with the counselor's speculations about the variance of individual treatment responses and their distribution, the counselor can use the findings of the present study to better understand the probabilities of various treatment outcomes.

With respect to the probabilities of a negative response to treatment, if the counselor is willing to assume that all clients have the exact same response to a treatment, then he or

she only need attend to the sample sizes and effect size. Generally speaking, if the effect size is at least 0.50 or the sample size is at least 15 in each group, the probability is less than 0.10 that the true effect size is negative. If clients do not have a uniform response to treatment, then when the effect size is less than the standard deviation of the individual responses to treatment, roughly 20% or more of the clients will have a negative response to treatment. This percentage will increase as the effect size decreases and the standard deviation of the individual responses to treatment increases, regardless of the sample size. All other things being equal, one can expect higher probabilities of a negative response if the distribution of individual responses is a mixture of extremely different subpopulations. In this latter instance, the results can be very different than what would be expected if there were a uniform response to treatment. In the worst case investigated, given only the sample size and effect size, one would conclude that the chances of the true effect size being negative was only 1 in 50, but if the standard deviation of individual responses and the shape of the distribution were taken into account, 56% of clients would be expected to have a negative response.

If a counselor is willing to assume that all clients have the exact same response to a treatment, then he or she only need attend to the sample sizes and effect size when determining the probability that the true effect size exceeds the minimal acceptable benefit of one half standard deviation increase on the outcome measure. Whenever the effect size exceeds 0.50, the probability that the true value exceeds 0.50 will be greater than 0.50. This probability increases as the sample size or the effect size increase. If clients vary in their response to treatment, this variability and the distribution of the individual treatment responses will combine, and depending on the specific conditions, may result in probabilities higher or lower than what would be expected if the clients responded uniformly. For symmetric distributions of individual responses to treatment, if the effect size equals 0.50, then the variability of individual responses has no effect, and the results are the same as if clients had a uniform response. Positively skewed distributions give lower probabilities, while negatively skewed distributions result in higher ones. If the effect size is 0.80, then all levels of variability and all distributions lead to probabilities that are *less than* would be expected if the clients all responded in the same manner. When the effect size is 0.30, with only three exceptions out of 45 sets of conditions, the probabilities are *higher than* would be expected if the clients all had the same response to treatment.

The utility of the results presented here depends on the realism of the models investigated. It is an open question as to how often one might encounter the variances and distributions studied. Perhaps the most that can be expected is that these results will serve to sketch the boundaries of an area with which counselors should be concerned. Until counselors can adequately estimate or control treatment reaction variability, much of the counseling outcome research reported will provide little justification for their interventions with clients.

References

Fisher, R.A. (1950). The fiducial argument in statistical inference. Contributions to Mathematical Statistics, pp. 390-398. New York: John Wiley & Sons, Inc. (Reprinted from Annals of Eugenics, Vol. VI, Pt. IV, pp. 391-398, 1935.)

Glass, G. V., & Hopkins, K. D. (1984). Statistical methods in education and psychology, 2nd ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. San Diego, CA: Academic Press, Inc.

Johnson, N. L., & Kotz, S. (1970). Continuous univariate distributions-1. New York: John Wiley & Sons, Inc.

Johnson, N. L., & Kotz, S. (1970). Continuous univariate distributions-2. New York: John Wiley & Sons, Inc.

MathSoft, Inc. (1993.) Mathcad 4.0 [Computer program]. Cambridge, MA.

Wood, P. K., & Games, P. (1990). Rationale, detection, and implications of interactions between independent variables and unmeasured variables in linear models. Multivariate Behavioral Research, 25(3), 295-311.

Zelen, M., & Severo, N. C. (1970). Probability functions. In M. Abramowitz & I. A. Stegun (Eds.), Handbook of mathematical functions (pp. 925-996). New York: Dover Publications, Inc.

Notes

Note 1:

The exact variance of ES, σ^2_{ES}, is given by Hedges and Olkin, 1985, p. 104 and differs from the definition used here, namely, $\frac{2}{n}$. For the values of μ_{ES} and n used here, the exact variance differs by no more than several hundredths, and this was considered inconsequential for present purposes. The definition of the variance used here assumes that the treatment and control populations have the same variance and that the "pooled" estimate of the variance is used to obtain the standard deviation in the denominator of ES. For the model defined above, whether or not the treatment and control populations have the same variance depends on the size of $\sigma_{\alpha_{Tj}}$ and the covariance of α_{Tj} and ε_{Tj}.

Negative values of the covariance can lead to the treatment population having a smaller variance than that of the control. Increasing the variance has the same effect as decreasing the sample size. Therefore, some insight into the effect of a larger treatment variance can be gained from looking at results for smaller sample sizes. However, as the results demonstrate, sample size has little effect on the probabilities reported, and therefore unequal variances would not be expected to affect the results either.

Note 2:

Specifying a likelihood function for μ_{ES} amounts to treating a parameter as a random variable. While there are different justifications for this, the one that seems most appropriate for present purposes was given by Fisher(1935) in an exposition of fiducial probability. Using the normal distribution to define the likelihood function for μ_{ES} finds some support in Hedges and Olkin's(1985) work where they construct confidence intervals for ES based on the normal distribution when sample sizes are of the size studied here. They actually used an unbiased estimator defined differently than that used here, but for the degrees of freedom studied here, the amount of bias is small. To the extent that the normal distribution suffers as an approximation due to its having a different variance than the exact distribution, this should not be of concern, for the results show that varying n, and, consequently, variance, has very little affect on the results when σ_{IES} is in the range investigated here.

Appendix

Table 1
Probabilities of a Negative Response

n	ES	σ_{ES}	Normal	Uniform	Chi2 +	Chi2 -	Mixture	Reference
15	.30	.33	.27	.28	.28	.25	.29	
15	.30	.67	.35	.37	.39	.29	.43	.21
15	.30	1.00	.39	.41	.45	.32	.49	
15	.50	.33	.16	.16	.15	.15	.17	
15	.50	.67	.26	.29	.27	.22	.33	.09
15	.50	1.00	.32	.36	.36	.26	.43	
15	.80	.33	.05	.05	.04	.06	.08	
15	.80	.67	.15	.17	.12	.14	.21	.01
15	.80	1.00	.23	.27	.23	.19	.32	
35	.30	.33	.23	.25	.23	.21	.28	
35	.30	.67	.34	.37	.39	.28	.46	.10
35	.30	1.00	.39	.41	.46	.31	.52	
35	.50	.33	.11	.12	.09	.11	.16	
35	.50	.67	.24	.28	.25	.20	.35	.02
35	.50	1.00	.31	.36	.37	.25	.45	
35	.80	.33	.03	.02	.01	.04	.08	
35	.80	.67	.13	.16	.08	.13	.22	.00
35	.80	1.00	.22	.27	.22	.19	.34	
90	.30	.33	.21	.24	.20	.18	.30	
90	.30	.67	.33	.37	.39	.27	.49	.02
90	.30	1.00	.38	.41	.46	.31	.56	
90	.50	.33	.09	.09	.04	.09	.17	
90	.50	.67	.23	.28	.24	.20	.38	.00
90	.50	1.00	.31	.36	.37	.25	.48	
90	.80	.33	.01	.00	.00	.03	.08	
90	.80	.67	.12	.15	.06	.12	.24	.00
90	.80	1.00	.21	.27	.22	.18	.37	

Table 2
Probabilities of an Acceptable Benefit

n	ES	σ_{ES}	Normal	Uniform	Chi2 +	Chi2 -	Mixture	Reference
15	.30	.33	.34	.35	.32	.36	.36	
15	.30	.67	.40	.41	.34	.45	.27	.29
15	.30	1.00	.43	.44	.35	.50	.17	
15	.50	.33	.50	.50	.47	.53	.54	
15	.50	.67	.50	.50	.44	.56	.44	.50
15	.50	1.00	.50	.50	.42	.58	.28	
15	.80	.33	.73	.72	.72	.75	.77	
15	.80	.67	.65	.63	.61	.71	.64	.79
15	.80	1.00	.61	.59	.55	.68	.46	
35	.30	.33	.31	.33	.27	.33	.37	
35	.30	.67	.39	.41	.32	.45	.22	.20
35	.30	1.00	.42	.44	.34	.51	.12	
35	.50	.33	.50	.50	.45	.55	.59	
35	.50	.67	.50	.50	.42	.58	.41	.50
35	.50	1.00	.50	.50	.41	.59	.24	
35	.80	.33	.77	.75	.77	.79	.84	
35	.80	.67	.66	.63	.61	.72	.62	.90
35	.80	1.00	.61	.59	.54	.69	.43	
90	.30	.33	.29	.33	.24	.32	.41	
90	.30	.67	.38	.41	.31	.46	.22	.09
90	.30	1.00	.42	.44	.34	.51	.11	
90	.50	.33	.50	.50	.43	.57	.67	
90	.50	.67	.50	.50	.41	.59	.41	.50
90	.50	1.00	.50	.50	.41	.59	.23	
90	.80	.33	.79	.76	.80	.82	.87	
90	.80	.67	.67	.63	.61	.73	.62	.98
90	.80	1.00	.62	.59	.54	.69	.42	